64392

Sensitivity Analysis for Hidden Bias in the Risk Assessments Conducted for the Federal Highway Administration Vision Waiver Program

FHWA-99-5578-6

September 1999

Sensitivity Analysis for Hidden Bias in the Risk Assessments Conducted for the Federal Highway Administration Vision Waiver Program

Risk assessment in the Federal Highway Administration's (FHWA) Vision Waiver Program was conducted through monitoring the accident experience of commercial motor vehicle (CMV) drivers who received waivers. Throughout the program, the accident rates of these drivers were compared to an accepted measure of the national accident rate for the population of this type of driver. In the assessment, a group receiving a treatment (i.e., being in the Waiver Program) was compared to a population of controls. This type of comparative design is common in studies, for example, in occupational health where disease and mortality rates in a sample of exposed workers are compared to rates for the nation as a whole representing the national norm (e.g., see Nicholson, Selikoff, Serdman, Lilis and Formby, 1979 - asbestos miners in Quebec). The strength of this design is it's apparent external validity with the national norm as the focus of comparison. However, the approach is not free of criticism and must be addressed to determine the overall value of the results.

The assessment of risk in the Vision Waiver Program has a research design which places it in the general category of observational studies. An observational study is an empirical investigation of treatments, policies, or exposure and the effects they cause. It differs from an experiment in that the investigation cannot control the assignment of treatments to subjects and therefore, cannot control the possibility of comparison bias because the sameness of those being compared cannot be guaranteed. There are two types of comparison bias; overt and hidden bias. Overt bias is related to the ancillary variables (covariates) that are measured and used in the study. For example, when collecting data about risk and exposure, other data can be obtained describing such covariates as an individual's gender, marital status, age, and work history. Using these variables, the possibility of overt bias can be investigated and, if it is present, these variables can be used to make adjustments. Hidden bias occurs when the variables that can cause bias are not measured and therefore, cannot be investigated in a straight forward manner. An example of this would be if age were strongly related to the cause of risk, but was not obtained for individuals in the study. Thus, without these data, the degree of potential bias could not be investigated and would therefore, be hidden.

The criticism of the risk assessment performed in the Vision Waiver Program was concerned with the issues involving comparison bias just mentioned. In the Waiver Program, risk assessment was performed by comparing the waivered drivers to the national accident rates in the National Highway Traffic Safety Administration's (NHTSA) General Estimates System (GES; see Appendix A for a note on GES). Some program reviewers argued that the comparison was a weak paradigm. They said that comparison bias could not be investigated because covariates were not available. They were right, the possibility of bias being present was hidden. The present investigation is designed to address that criticism and to determine if the comparison is vulnerable to hidden bias.

The approach taken in this investigation is a sensitivity analysis concerning the possible lack of balance among unmeasured covariates in the waiver group and the GES data. By contrast, in an experimental study, where subjects are randomly assigned to groups, balance is guaranteed, at least theoretically. In observational studies, such as the risk assessment, there can be a lack of covariate balance between the groups compared. If the lack of balance is among covariates that are related to the outcome (i.e., accidents in this case), comparison bias is a possibility. When bias is hidden, investigating this possibility using sensitivity analysis essentially involves examining how various degrees of imbalance affect the outcomes.

The sensitivity analysis in this investigation addresses two facets of risk assessment in the Waiver Project. One is the last monitoring comparison with GES, conducted while the program was in effect. The other was an assessment for the nine months after the program was closed as a follow-up assessment of risk (i.e., the last 9 months of 1996). Performing sensitivity analyses on the two facets of the program essentially covers all risk assessment performed. Thus, there is ample opportunity to establish the validity of the risk assessments.

Model for Sensitivity Analysis

To conduct a sensitivity analysis for hidden bias, a framework, or model, is needed through which the analysis can be systematically performed. In general, this involves the use of a statistical form which can embody the comparison of the exposed sample (i.e., the Vision Waiver group) to the national data. In addition, there also is the need to have a medium through which the sensitivity analysis can be structured. This medium should allow the range of values tested in the sensitivity analysis to be entered into the above mentioned statistical form. In the case of comparing accident experience, the statistical form to be used is given as:

$$z = (n - E) / \mathrm{sqrt}[\mathrm{Var}(E)]$$

where

n = The number of accidents experienced by the waiver group in a specific period of time.

E = The expected number of accidents in the period based the national accident rate (r) and the exposure of the waiver group in vehicle miles traveled (VMT) (i.e., E = VMTr).

sqrt[Var(E)] = The square root of the variance of E.

z = A variate with a standard normal probability distribution

The use of multiple values in the sensitivity analysis is carried out through the national accident rate and the degree to which it is believed that the unmeasured variables in the comparison are unbalanced. The medium for combining these factors is:

$$v = (L * p) / [(1 - p) + (L * p)]$$

where

p = The probability of one or more accidents in the national truck driver population.
= 1 - exp(-r) [exp is the natural base and (-r) is it's exponent]

L = The degree of imbalance in unmeasured variables between the waiver group and the national data (L can equal 1, 2, 3, etc. or 1/2, 1/3, 1/4, etc., for balance in both directions).

v = The adjusted probability of one or more accidents in the national population - used to convert back to an adjusted national accident rate.

The sensitivity analysis is carried out by varying values of L (i.e., degree of imbalance) and determining the effect on the statistical significance of z. If the significance level changes at some value of L, that is used to determine how sensitive, or insensitive, the results are to hidden bias. Should no changes occur until L = 3 or 4 or 1/3 or 1/4, one could say that the results are relatively insensitive to hidden bias. The larger the values of L become before changes in significance occur, the more robust the original results can be viewed. The approach to sensitivity analyses is given by Rosenbaum (1995) and is explained more fully in Appendix B.

Results

Sensitivity analyses were performed, as was mentioned earlier, for two facets of the Waiver Program. The first tested the results of the last risk assessment performed when the program was in effect. The second analysis examined an assessment conducted for the nine months following the termination of the program.

Sensitivity Analysis for the Risk Assessment Performed During the Program

The results of the sensitivity analysis for the risk assessment performed during the Waiver Program are given in Table 1. The assessment was conducted in November 1995 for the 2,237 drivers that were in the program at that time. The number of accidents and VMT used in calculating the accident rate for the waiver group were those experienced from when they entered the program until November 1995 (i.e., 528 accidents and 300.15 million VMT - these are slightly different from those in the November 1995 Monitoring Report since they were updated based on continued surveillance). The accident rate for these drivers was 1.759 accidents per million VMT.

The national accident rate in Table 1 is given as 2.348 based on GES data combined from 1992 through 1995. When this is compared to the waiver group's rate using the statistic explained earlier, the resulting z statistic is -6.677. This result is insignificant in this context because the concern only involves whether the waiver group's rate is larger than the national expectation which, of course, it is not.

The results of the sensitivity analysis for the comparison are given in the lower portion of Table 1. The values of imbalance among unmeasured covariates (L) is given next to the effect they have on the value of the z statistic. In the lower left columns where the values of L range from 1 (no imbalance) to 3, no change in the significance of z occurs. If the sensitivity analysis made the rate of the waiver group significantly larger than the national rate, the value of z would be larger than 1.65.

The columns on the right present the degrees of imbalance in the other direction. Here it is seen that when unmeasured variables have a one-third imbalance relative to being in the waiver group (i.e., L = 1/3), the effect on z is a change in significance (i.e., 4.899 is larger than 1.65). At this point, the results are potentially sensitive to hidden bias.

Sensitivity Analysis for the Risk Assessment After the Program Ended (Apr. - Dec. 1996)

The results of this sensitivity analysis are given in Table 2. In the period after the Waiver Program terminated that is the focus of this analysis, the waiver group experienced 59 accidents with 71.97 million VMT. This gives an accident rate of 0.820 accidents per million VMT. In the calendar year 1996, the national rate, given by GES, was 2.068. The z statistic for the waiver group based on the national expectation was -7.364 (i.e., when L = 1).

As can be seen by the results in the lower part of the table, none of the values of L presented produced a value for z that would claim that the rate of the waiver group is significantly larger than the national rate (i.e., z = 1.65 for significance at the .05 level for a one-sided test). In fact, L would need a value of less than 1/7 before the significance of the original test is altered.

Conclusions

These risk assessments have some insensitivity to hidden bias. The assessment conducted during the program appears to be less so than that done after the program closed. The assessment performed after the termination of the Waiver Program is apparently very sensitive to hidden bias. As was shown, the results could be reversed if a hidden factor critical to accidents were under represented in the waiver group by a factor of seven. Even in an observational study this is highly unlikely.

The last risk assessment conducted during the program could also be considered somewhat insensitive to hidden bias. There, some unknown critical characteristic or combination of characteristics would have to be under represented by one-third in the waiver group before their risk exceeded that of the national norm. On one hand, it should be somewhat perplexing to have the sensitivity to hidden bias change to some degree when the analysis was performed on basically the same set of drivers. The only facet of the investigation that clearly changed is time.

However, on the other hand, this may give some insight to explaining the increased insensitivity to bias in the post-program monitoring.

It must be recalled that sensitivity analysis is performed because information is not available to investigate bias in an overt fashion. Thus, a change in sensitivities for the same group across time must, of course, be approached with caution. But, even with this in mind, there is some evidence which offers a plausible explanation. It has to do with age of the waiver group. In 1995, the median age of those still in the waiver group was 48. This is the group that was to receive grandfather rights in March of 1996 and the group upon which the second sensitivity analysis was performed.

In 1996, a very large random sample of driving records was obtained from the Commercial Driver's License Information System (CDLIS). The sample was in excess of 110,000 records which suggests very little sampling error. The median age of this group was 35 in 1995, a difference of 13 years younger than the waiver group. Since accident rates have often been linked with age, this evidence along with some others could explain changes in bias sensitivity and differences in accident rates.

The first sensitivity analysis was performed on the group as it was exposed to four years of driving (1992 to 1995). This period would allow both groups (waiver and national population) to be operating CMVs at younger ages. However, since the waiver group is a much smaller subset (2,176) of the national CMV operator population, it should be more sensitive to changes in accident rate. This, in fact, is the case. In a comprehensive risk assessment perforrned on the Vision Waiver Program before its members received grandfather rights, an examination of the periodic accident rates showed a significant decrease across time for the group that survived to receive those rights.

A possible interpretation could then be related to both possible age differences and the period of comparison. When both groups were younger and observed over a longer period, age could have had more of an influence as a hidden biasing factor (i.e., a reduced insensitivity to bias). Whereas, when the groups were both older and compared for barely a year, the hidden biasing influence of age would be less severe (i.e., high insensitivity to hidden bias).

In this sense, age could be an explanation for the differences in the results of the sensitivity analyses. While this seems a plausible explanation of comparison sensitivities, it is nonetheless, without direct (i.e., overt) data for adjustment. Other factors could play a role in the form of the results. However, even the sensitivity analysis for the longer period showed that the results were relatively insensitive to hidden bias and one could argue that the results of the original risk assessments have reasonable validity.

References

Nicholson, W. et al (1979) "Long term mortality experience of chrysotile miners and millers in Thetford Mines, Quebec", Annuals of the New York Academy of Sciences, 330, 11-21.

Rosenbaum, P.R. (1995) Observational Studies, Springer-Verlag, New York.

Table 1. Results of the Sensitivity Analysis for Hidden Bias in the Comparison of the Accident Rate for the Vision Waiver Group to the National Accident Rate (1992-1995).

- Results from comparison in monitoring report (November 1995, 2,237 waivered drivers)

Waiver Group's Rate[1]	National Rate[2]
1.759	2.348

- Results of sensitivity analysis

L	z	L	
1	-6.677		
2	-12.328	1/2	0.204
3	-15.260	1/3	4.837***

---

[1]Rate based on 528 accidents and 300.15 million VMT by the waiver group between July 1992 and November 1995.

[2]National accident rate is given for large trucks by the General Estimates System (GES, see Appendix A.). Rate is based on an estimated 1,553,607 accidents for 1992 through 1995 and 661,644 million VMT in that period.

***Significant at $p < .01$.

Table 2. Results of the Sensitivity Analysis for Hidden Bias in the Comparison of the Accident Rate for the Vision Waiver Group to the National Accident Rate After the Program had Terminated (April 1996 to December 1996).

- Results for a comparison after the Waiver Program had terminated (2,176 grandfathered drivers)

Waiver Group's Rate[1]	National Rate[2]
0.820	2.068

- Results of sensitivity analysis

L	z	L	z
1	-7.364	1/2	-4.696
2	-9.707	1/3	-2.929
3	-10.933	1/4	-1.573
4	-11.752	1/5	-0.458

---

[1]Rate based on 59 accidents and 71.97 million VMT in the April - December 1996 period.

[2]National accident rate is based on 378,335 accidents given to GES (see Appendix A) and 182,97 1 million VMT.

Appendix A

A Note on the General Estimates System (GES)

Appendix A

A Note on the General Estimates System (GES)

The General Estimates System (GES) was selected as the best measure of the prevailing national norm relative to large truck accidents. The GES is operated by the National Center for Statistics and Analysis (NCSA), an office within the National Highway Traffic Safety Administration (NHTSA). GES is a survey of police accident reports (PARs) in the United States.

The GES is based on a probability sample of about 45,000 PARs selected on an annual basis. The selection of PARs is accomplished in three stages. The first stage is a sample of geographic areas from across the United States. These Primary Sampling Units (PSUs) were formed by the University of Michigan Transportation Institute and could be a central city, a county surrounding a central city or a group of contiguous counties. The second stage in the sample selection process is a sample of police jurisdictions (PJ) within the PSUs. The third stage of sampling involves the selection of PARs within a PJ. (See NHTSA Technical Rep. DOT/HS/807-796; Dec. 199 1.)

Since 1988, GES has over-sampled accidents involving large trucks. In the 1992 sample, 6861 large truck accidents are included in the data. Truck accidents are included in this sample if the vehicle is a medium (10,001 to 26,000 lbs) or a heavy (over 26,000 lbs) truck and at least one passenger vehicle was towed or an involved person was injured.

In 1992, data from GES provided an estimate of 386,000 accidents involving a large truck. Using the estimated total number of accidents for large trucks, a national accident rate for this class of truck can be found. The Office of Highway Information Management (OHIM) of the

Federal Highway Administration (FHWA) has calculated that large trucks had 152,538 million vehicle miles traveled (VMT) in the United States in 1992. A national accident rate for large trucks in 1992 would be obtained as 386,000/1 52,538 or 2.53 1 accidents per million VMT.

The national accident rate and the number of accidents in the sample can be used to estimate the extent of coverage the GES provides as a reference group for the waivered drivers. That is, the expected amount of VMT associated with the 6,861 truck accidents is obtained as 6,861/2.53 1 or 2,710.8 million VMT exposure. This is the expected VMT exposure needed to generate the number of accidents in the sample. If large truck drivers drive on the average of 40,000 miles per year, then the sample of 6,861 accidents represents more than 67,000 individuals (2,710,800,000/40,000 = 67,700).

To some degree, the national accident rate, and its apparent large sample, makes the study similar to a decision problem with a fixed criterion. In that circumstance, the criterion (e.g., a national accident rate) would have no sampling variance. The conservative decision would then show the waiver group as having acceptable safe driving behavior if their rate was less than or equal to the national rate. The question of the waiver rate being above the national rate would never arise. The only decision considered would be concerned with how confident the decision makers wish to be that the sample accident rate of the waiver group is actually less than or equal to the national criterion given the sampling error in the waiver group.

This type of decision framework is not quite available in the present circumstance. However, with an apparent large national sample, there is no reason why the national rate cannot be treated as a constant. In this case, only the waiver group rate would have sampling variance. Here, then, the sample rate of the waiver group could only exceed the national rate by a small amount before being judged significantly too large. This places more constraint on the waiver group's accident rate than would be provided by a smaller control group. Moreover, the national rate, as a comparison, enhances the element of external validity.

Appendix B

Sensitivity Analysis for Hidden Bias

Appendix B

Sensitivity Analysis for Hidden Bias

If a study is free of hidden bias, then the chance that an unit of analysis receives the treatment (exposure) is some function of the observed covariates describing the unit. There is hidden bias if two units with the same values for the covariates have differing chances of receiving treatments. Here, units that appear to be similar actually have different probabilities of assignment to treatments. The overall approach to sensitivity analyzed for hidden bias is given in Rosenbaum (1995).

In cases like the latter above, a sensitivity analysis examines how the results concerning treatment effects would be altered by hidden biases of varying degrees of magnitude. That is, given the chances of assignment to treatment differ for the same covariate values, how large does that difference have to be to alter the inferential conclusions of the study? Stated differently, how unbalanced do the covariates have to be, relative to treatment, before the significance of the results changes from the original unexamined findings?

The lack of balance among covariates (unmeasured), can be represented as an odds ratio. If, for example, the probability of one unit being assigned to treatment is given as s, then the odds that the unit receives treatment is:

$$s / (1 - s)$$

Similarly, if another unit with the same covariate values has a probability of q (s not = q) of receiving treatment, then the odds of this unit receiving treatment is:

$$q / (1 - q)$$

For units with the same covariate values, the odds ratio for assignment to treatment is:

$$0 = [s * (1 - q)] / [q * (1 - s)]$$

If o = 1, the two units have the same chances of being assigned to the treatment (i.e., s = q). However, if o is less than or greater than 1, it is a measure of the lack of balance in the covariates relative to treatment assignment. In these circumstances, the odds ratio (o) can have the value L (or l/L) where L can have degrees of imbalance greater (less) than 1 as L = 1, 2, 3, 4, etc.

Sensitivity analysis, using L, can be integrated into study results in numerous ways depending on the type of data involved and the statistical methods used. For the risk assessment performed in the Vision Waiver Program, the following type of approach was used:

$$v = (L * p) / [(1 - p) - (L * p)]$$

where

L = The value of the odds ratio explained above when it departs from 1.

p = The expected probability of an adverse outcome in the comparison group as original unadjusted data (e.g., the accident rate in GES data for a given set of years).

v = The adjusted probability of an adverse outcome (i.e., adjusted for covariate imbalance through L).

To see how this framework fits into the present risk assessment, data from that work will be used as an example. In the last risk assessment performed in the Waiver Program (November 1995), the accident rate of the Vision Waiver group was compared to the national accident rate given by GES (see Appendix A). Given in the updated seventh monitoring report for the program, an accident rate of 1.759 for the waiver group was compared to a national (GES) rate of 2.348. The waiver group's rate was based on 528 accidents occurring during 300.15 million vehicle miles traveled (VMT). In the absence of hidden bias, the normal test of no treatment effect considers the 528 accidents as a poisson count with expectation of 7 15.80 (from the GES data). The test applied to these data is a z statistic given as:

$$z = [528 - (300.15) * (2.348)] / \mathrm{sqrt}[(300.15) * (2.348)]$$

where sqrt [ ] is the square root operator and (300.15) * (2.348) is the variance of the expected value. A one-sided z test is used in this circumstance because the only concern was whether the waiver group's accident rate significantly exceeded the GES rate. In the above test, assuming no hidden bias, z = -6.658 which is considerably removed from the z value of 1.64 (type 1 error = .05) above which the waiver group's accident rate would be significantly larger than the national (GES) rate.

The sensitivity analysis for hidden bias is carried out through the national (GES) accident rate. This is done using the adjusted values for v generated by examining different levels of imbalance, in covariates through L. In particular, in the workup for v, the value of p represents the probability of one or more accidents and is given by:

$$p = 1 - \exp(-2.348)$$

where exp(t) represents the natural base(e) raised to the power oft. In this case, p = .904. The adjustment for v is then given as:

$$v = [(.904) * L] / [(1 - .904) + (.904 * L)]$$

with L = 1, 2, 3, etc. or 1/2, 1/3, 1/4, etc. Since the sensitivity analysis is conducted through changes in the national rate (based on L), that rate is recaptured through v as ln(1 - v) = adjusted national rate to be used in the z statistic. This procedure generated the findings in the results section.